What is claimed is:

1. A method of classifying a thing as a member of one or more out of a plurality of classes, said thing having a plurality of attributes associated therewith, said method comprising the steps of:

(a) for each of said plurality of classes, assigning attribute values based on each of said attributes, each said attribute value representative of a relative possibility that said thing is a member of the associated class based on said attribute,

(b) for each of said plurality of classes, aggregating said attribute values using a t-norm function,

(c) selecting a highest aggregated value,

(d) determining that said thing belongs to the class associated with said highest aggregated value, and

(e) determining a confidence factor based on the relative magnitude of said highest aggregated value and a second highest aggregated value.

2. The method of claim 1 further comprising:

(f) normalizing said attribute values based on the relative information provided by each attribute.

3. A method of training a machine to classify a thing as a member of one or more out of a plurality of classes, the method comprising the steps of:

(a) providing training data to said machine, said training data comprising a plurality of records, each record having attribute data associated therewith, said attribute

data comprising values associated with a plurality of possible attributes, each record further having a class value associated therewith indicating the class to which the record belongs,

(b) for each of said possible attributes, normalizing said attribute data for each record based on the distribution of values present for the attribute in substantially all of said records,

(c) for each of said records, performing a t-norm operation on the available attribute data, and generating a possibility value for each of said possible classes, said possibility values corresponding to the relative possibility that the record belongs to one of said particular classes,

(d) for each of said plurality of classes, aggregating substantially all of the records having the class value associated with said class, and generating weights for each of the attributes according to the degree that each attribute corresponds with a correct determination of said class.

4. The method of claim 3, further comprising the steps of:

(e) for each of said records, generating belief values for the one or more classes having the highest possibility values, said belief value representing the difference between the possibility value for said class, and the next highest possibility value, and

(f) generating a list of informative attributes from the attributes associated with records for which belief values above a threshold value were generated.

5. An article of manufacture adapted to be used by a computer, comprising:

a memory medium on which are stored machine instructions that implement a plurality of functions useful for classifying an item as a member of one or more out of a plurality of classes, said thing having a plurality of attributes associated therewith, when the machine instructions are executed by a computer, said functions including:

(a) for each of said plurality of classes, assigning attribute values based on each of said attributes, each said attribute value representative of a relative probability that said thing is a member of the associated class based on said attribute,

(b) for each of said plurality of classes, aggregating said attribute values using a t-norm function,

(c) selecting a highest aggregated value,

(d) determining that said thing belongs to the class associated with said highest aggregated value, and

(e) determining a confidence factor based on the relative magnitude of said highest aggregated value and a second highest aggregated value.

6. An article of manufacture adapted to be used by a computer, comprising:

a memory medium on which are stored machine instructions that implement a plurality of functions useful for training a machine to classify a thing as a member of one or more out of a plurality of classes, said functions including:

(a) providing training data to said computer, said training data comprising a plurality of records, each record having attribute data associated therewith, said attribute data comprising values associated with a plurality of possible attributes, each record further having a class value associated therewith indicating the class to which the record belongs,

(b) for each of said possible attributes, normalizing said attribute data for each record based on the distribution of values present for the attribute in substantially all of said records,

(c) for each of said records, performing a t-norm operation on the available attribute data, and generating a possibility value for each of said possible classes, said possibility values corresponding to the relative possibility that the record belongs to one of said particular classes,

(d) for each of said plurality of classes, aggregating substantially all of the records having the class value associated with said class, and generating weights for each of the attributes according to the degree that each attribute corresponds with a correct determination of said class.

7. The article of claim 6, said functions further including:

(e) for each of said records, generating belief values for the one or more classes having the highest possibility values, said belief value representing the difference between the possibility value for said class, and the next highest possibility value, and

(f) generating a list of informative attributes from the attributes associated with records for which belief values above a threshold value were generated.

8. An apparatus adapted to classify a thing as a member of one or more out of a plurality of classes, said thing having a plurality of attributes associated therewith, said apparatus comprising:

an output device and an input device,

a processor, and

a memory having machine executable instructions for performing a series of functions stored therein, and adapted to receive and store a series of data records, said functions including:

(a) receiving at said input device a data record corresponding to said thing sought to be classified, said data record comprising attribute values corresponding to the attributes of said thing,

(b) for each of said plurality of classes, generating an aggregated value by aggregating said attribute values using a t-norm function,

(c) selecting a highest aggregated value from said aggregated values,

(d) determining a most possible class from among the plurality of classes based on said highest aggregated value,

(e) determining a confidence factor based on the relative magnitude of said highest aggregated value and a second highest aggregated value, and

(f) outputting said most possible class and said confidence factor at said output device.

9. An apparatus adapted to be trained to classify a thing as a member of one or more out of a plurality of classes, said thing having a plurality of attributes associated therewith, said machine comprising:

an output device and an input device,

a processor, and

a memory having machine executable instructions for performing a series of functions stored therein, and adapted to receive and store a series of data records, said functions including:

(a) receiving training data at said input device, said training data comprising a plurality of records, each record having attribute data associated therewith, said attribute data comprising values associated with a plurality of attributes, each record further having a class value associated therewith indicating the class to which the record belongs,

(b) for each of said attributes, normalizing said attribute data for each record based on the distribution of values present for the attribute in substantially all of said records,

(c) for each of said records, performing a t-norm operation on the available attribute data, and generating a possibility value for each of said possible classes, said possibility values corresponding to the relative possibility that the record belongs to one of said particular classes,

(d) for each of said plurality of classes, aggregating substantially all of the records having the class value associated with said class, and generating weights for each of the attributes according to the degree that each attribute corresponds with a correct determination of said class.

10. The apparatus of claim 9, said functions further comprising:

(e) for each of said records, generating belief values for the one or more classes having the highest possibility values, said belief value representing the difference between the possibility value for said class, and the next highest possibility value, and

(f) generating a list of informative attributes from the attributes associated with records for which belief values above a threshold value were generated.

11. The apparatus of claim 10, said functions further comprising:

(g) outputting said belief values and said list through said output device.

12. A neural network comprising:

at least an input layer and an output layer, the input layer having a plurality of input nodes, and the output layer having a plurality of output nodes, such that each of the output nodes receives weighted input from each of the input nodes representative of the possibility that the particular output node represents the correct output,

wherein the output nodes aggregate the input from each of the input nodes according to a t-norm function, and produce an output representative of the result of the t-norm function.

13. A neural network comprising:

at least an input layer, an output layer, and at least one confidence factor node, the input layer having a plurality of input nodes, and the output layer having a plurality of output nodes, such that each of the output nodes receives weighted input from each of the input nodes representative of the possibility that the particular output node represents the correct output, and the confidence factor node receives input from each of the output nodes,

wherein the output nodes aggregate the input from each of the input nodes according to a t-norm function, and produce an output representative of the result of the t-norm function, and wherein the confidence factor node produces an output representative of the difference between the highest output from the output nodes and the second highest output from the output nodes.

14. The neural network of claim 13 wherein the network includes a plurality of confidence factor nodes, each receiving input from each of the output nodes, and the output of each confidence factor node is representative of the difference between the output of the n highest output nodes and the next highest output from the output nodes.

15. A universal parallel distributed computation machine comprising:

at least an input layer and an output layer, said input layer having a plurality of input neurons, and said output layer having a plurality of output neurons, such that each of said neurons has a weight connection to at least one other neuron,

wherein said weight connection represents mutual information, and said mutual information is represented by a likelihood function of weight.

16. The machine of claim 15 wherein a value for said weight connections is determined by multiplying the likelihood functions for two associated neurons, and normalizing the result.

17. The machine of claim 15 wherein said machine is an analog parallel distributed machine.

18. The machine of claim 15 wherein said machine is a digital parallel distributed machine.

19. The machine of claim 15 wherein said machine is a hybrid digital and analog parallel distributed machine.

20. A method of training a neural network comprising an input layer having a plurality of input neurons and an output layer having a plurality of output neurons, each of said neurons having a weight connection to at least one other neuron, said method comprising the steps of:

(a) providing training data to said machine, said training data comprising a plurality of records, each record having at least one neuron associated therewith, such that said record causes said associated neuron to fire a signal to connected neurons,

(b) updating weights of said weight connections using a likelihood rule, said rule based on the likelihood of each connected neuron firing and of both neurons firing together,

(c) aggregating said signals at each said connected neuron with a t-conorm operation,

(d) evaluating the performance of said machine, and

(e) repeating steps (a)-(d).